Volume 15 Number 1 1987 Nucleic Acids Research

Assignment of the ^{13}C nuclear magnetic resonance spectrum of a short DNA-duplex with ^{1}H-detected two-dimensional heteronuclear correlation spectroscopy

Werner Leupin, Gerhard Wagner, William A.Denny[1] and Kurt Wüthrich

Institut für Molekularbiologie und Biophysik, ETH-Hönggerberg, CH-8093 Zürich, Switzerland and [1]Cancer Research Laboratory, University of Auckland School of Medicine, Private Bag, Auckland, New Zealand

Received November 4, 1986; Accepted December 1, 1986

ABSTRACT

Proton-detected ^{1}H-^{13}C heteronuclear correlated spectroscopy ([^{1}H,^{13}C]-COSY) was used to establish relations between the carbon-13 and proton nuclear magnetic resonance chemical shifts in the hexadeoxynucleoside pentaphosphate d-(GCATGC)$_2$. Using the previously established sequence-specific proton NMR assignments, sequence-specific assignments were thus obtained for nearly all proton-bearing carbons. This approach offers a new criterion for distinguishing between the proton NMR lines of purines and pyrimidines, based on the different proton-carbon-13 coupling constants. Furthermore, the adenine ring carbon 2 has a unique carbon-13 chemical shift, which enables a straightforward identification of the adenine C2H resonances by [^{1}H,^{13}C]-COSY.

INTRODUCTION

Most NMR studies of nucleotides have employed proton and phosphorous resonances, whereas ^{13}C and ^{15}N were little used because of the low sensitivity. With the introduction of heteronuclear 2D NMR experiments with proton detection (1-9), ^{13}C and ^{15}N resonances can now also be assigned conveniently and used to analyze nucleotides without isotope enrichment. In this report we demonstrate the use of heteronuclear correlated spectroscopy ([^{1}H,^{13}C]-COSY) with studies of the hexadeoxynucleoside pentaphosphate duplex d-(GCATGC)$_2$:

$$\begin{array}{l} d\text{-}{}^{5'}G_1\ C_2\ A_3\ T_4\ G_5\ C_6{}^{3'} \\ \qquad C_{-1}\ G_{-2}\ T_{-3}\ A_{-4}\ C_{-5}\ G_{-6}\text{-}d \end{array}$$

We use an experiment designed for efficient suppression of the signals from carbon-12 bound protons, enabling easy observation of the much weaker signals from the ^{13}C satellites (1,10,11). Previously, sequence-specific ^{13}C NMR assignments were obtained with [^{1}H,^{13}C]-COSY for some oligoribonucleo-

tides (12) and oligodeoxyribonucleotides (13) in single stranded form. For DNA fragments in duplex form (e.g., d-$(CGCGCG)_2$,d-$(CGCGCGCG)_2$; (14)) one-dimensional NMR methods were used and combined with empirical criteria for resonance assignments. The presently described two-dimensional NMR approach has the advantage of superior reliability and speed. It is of particular interest with regard to the potential of ^{13}C NMR spectroscopy for studies of structure and dynamics of oligonucleotides (15), as well as its use as a complementary method to ^{1}H NMR spectroscopy for studies of DNA-ligand interactions (e.g.,16).

MATERIALS AND METHODS

The hexadeoxynucleoside pentaphosphate d-GCATGC was synthesized by the phosphotriester method in liquid phase as described elsewhere (17,18). The oligonucleotide was used as the Na^+ salt (obtained by running the DNA over a Dowex 50W X8 column in the Na^+-form, with H_2O as the solvent, followed by desalting with Sephadex G-25). The NMR sample was prepared by dissolving the DNA in 500 µl buffer solution in D_2O, containing 0.1 M NaCl and 50 mM phosphate, p^2H = 7.0. This solution was twice lyophylized and redissolved in 500 µl of 99.96% D_2O. After a third lyophylization the final solvent used was 500 µl of 99.996% D_2O (Stohler Isotopes). The concentration of the DNA was 20 mM in duplex.

The NMR spectra were recorded in a 5 mm sample tube on a Bruker AM 360 spectrometer at 297±0.5 K. A [^{1}H,^{13}C]-COSY experiment with 490 t_1 values was recorded. The maximum t_1 value was 12.3 ms. Each free induction decay consisted of 1024 data points with a maximum t_2 value of 140 ms. Prior to Fourier transformation the time domain data set was multiplied with a trapezoidal window along t_2 , with the plateau reaching from 3 ms to 82 ms, and a cosine window was applied along ω_1. Chemical shifts are quoted relative to internal TSP (2,2,3,3-tetradeutero-3-(trimethylsilyl)propionate, sodium salt).

RESULTS AND DISCUSSION

Figure 1 shows a survey of the [^{1}H,^{13}C]-COSY spectrum of the hexadeoxynucleoside pentaphosphate d-(GCATGC) in duplex form. The ^{1}H spectrum is along ω_2 , the ^{13}C spectrum along ω_1. For each of the observed cross peaks the two doublet components have opposite sign. The assignments of the different groups of resonance peaks to particular CH or CH_2 fragments in the mononucleotides follows from the chemical shift ranges generally observed for the corresponding ^{13}C and ^{1}H lines (e.g.,19-22).

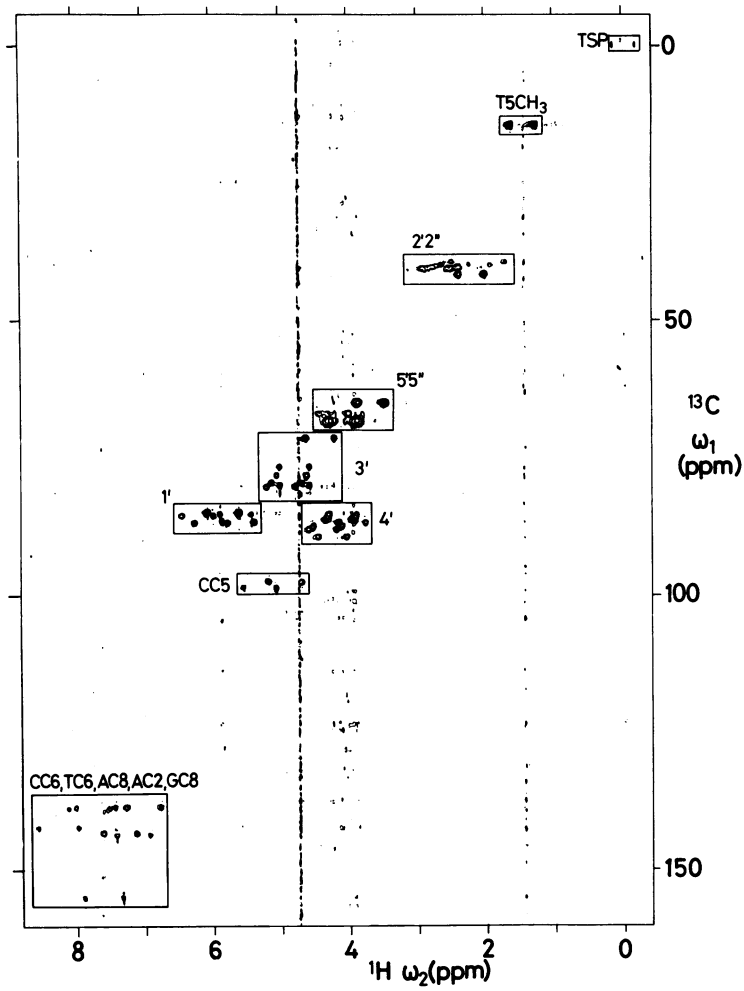
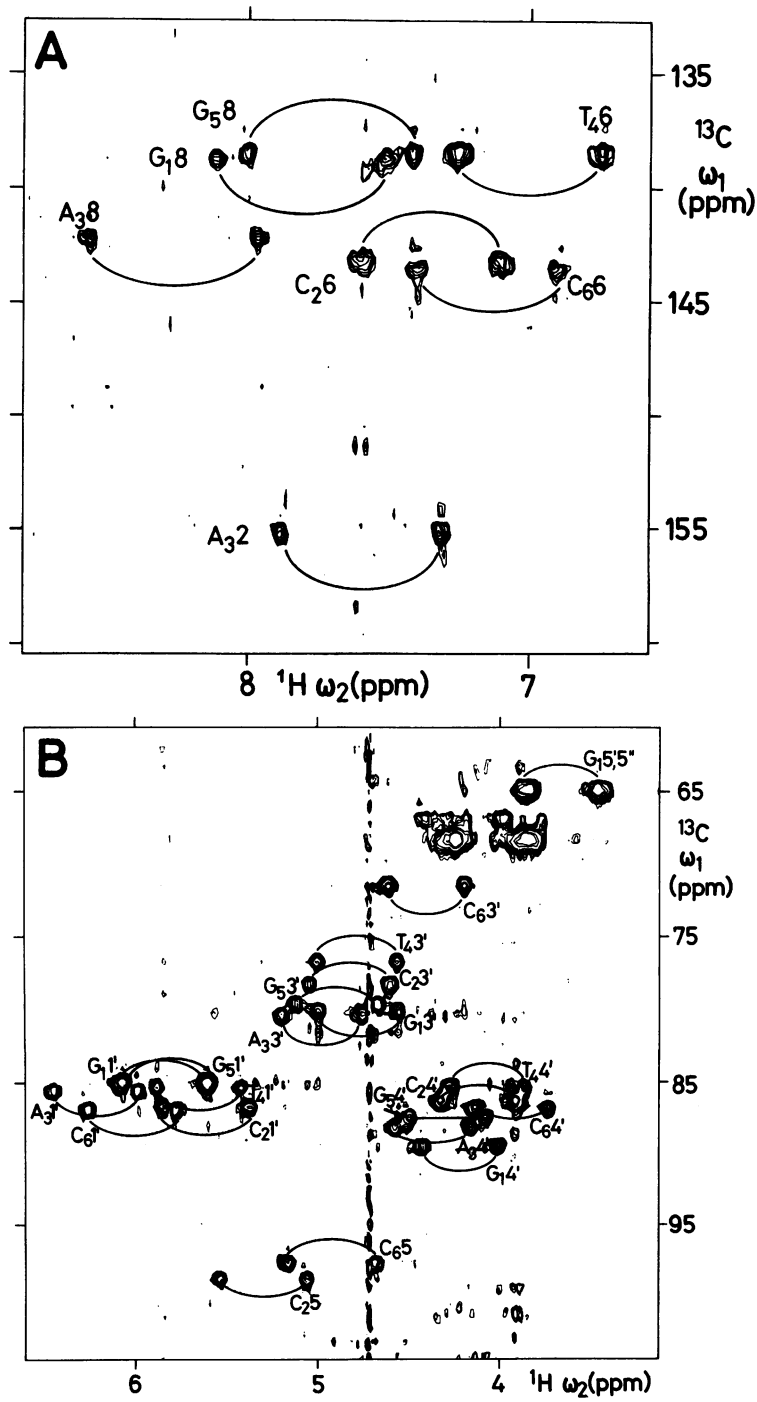


Figure 1: [$^{1}H,^{13}C$]-COSY of a 20 mM solution of d-(GCATGC)$_2$ in D_2O, containing 0.1 M NaCl, 0.05 M phosphate, pH=7.0, T=297 K. The total measuring time was 72 hours. The different types of cross peaks are boxed and labelled. TSP identifies the cross peak due to the internal reference. The cross peaks show an antiphase doublet splitting along ω_2 due to the heteronuclear coupling. The opposite signs of the two multiplet components (left component positive, right component negative) are not indicated in the figure.

Figure 2 contains expanded plots of three spectral regions in Fig.1. The multiplet splittings along ω_2, which are due to the ^{13}C-^{1}H couplings, are indicated with curved lines. Since the ^{1}H NMR spectrum of the hexanucleotide was previously assigned (17), sequence specific assignments for all ^{1}H-bea-


A
G5 8
G1 8
T4 6
A3 8
C2 6
C6 6
A3 2
135
145
155
13C ω1 (ppm)
8 1H ω2(ppm) 7
B
G1 5'5"
C6 3'
T4 3'
C2 3'
G5 3'
G1 3'
A3 3'
G1 1'
G5 1'
T4 1'
A3 1'
C6 1'
C2 1'
C2 4'
T4 4'
G5 4'
C6 4'
A3 4'
G1 4'
C6 5
C2 5
65
75
85
95
13C ω1 (ppm)
6 5 4 1H ω2(ppm)

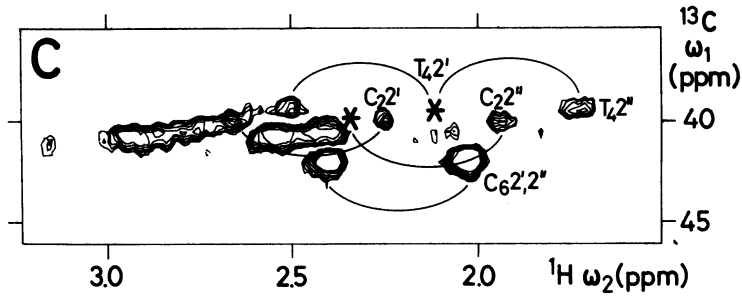


Figure 2: Expanded plots from the spectrum of Fig. 1. **(A)** Region of the cross peaks corresponding to ring C-H groups of the bases. **(B)** Region of the cross peaks corresponding to 1'CH, 3'CH, 4'CH, and 5'CH_2 fragments of the deoxyriboses. **(C)** Region of the cross peaks corresponding to the 2'CH_2 fragments of the deoxyriboses. The heteronuclear splittings of the cross peaks along ω_2 are indicated by curved lines. The peaks are identified by the one-letter abbreviations for the deoxynucleotides and the sequence position. In spectrum **(C)**, stars indicate the positions of multiplet components of cross peaks which were cancelled by overlap (see text).

ring ^{13}C-nuclei were obtained as indicated in the figure, the only exceptions being some C2' and C5' positions. Table 1 lists the chemical shifts and the heteronuclear coupling constants $^1J_{CH}$.

The sequence-specific ^{13}C resonance assignments were accomplished by a comparison of the ^1H chemical shifts in the spectrum of Fig.1 with the sequence-specific assignments derived earlier by ^1H NOESY and ^{31}P-relayed ^1H COSY (17). This was straightforward for all C-H fragments, but for the CH_2 groups only partial assignments of the carbon-13 resonances were obtained (Table 1). For all CH_2 groups limitations resulted from cancellation of cross peak components due to mutual overlap, which is especially pronounced in cases where the chemical shift dispersion of the ^{13}C and ^1H resonances is small, and for the 5'CH_2 groups we had further the situation that all 5'H and 5"H resonances had not been assigned (17). In the region of the [^1H,^{13}C]-COSY spectrum containing the cross peaks of the 2'CH_2-groups the cancellation of cross peak multiplet components due to two different reasons can be demonstrated: First, for T_4 the cancellation of two multiplet components is due to the fact that the positive left hand component of the cross peak between T_4C2' and $T_4$2'H overlaps with the negative right hand component of the T_4C2'-$T_4$2"H cross peak. This cancellation thus arises because the difference between the chemical shifts of $T_4$2'H and $T_4$2"H is equal to the coupling constants $^1J_{CH}$. Second, for C_2 the positive left hand component of the higher field cross peak C2'-2"H is cancelled due to overlap

Table 1: ^{13}C-NMR chemical shifts [a] and $^{1}J_{CH}$ [b] for the hexanucleotide d-(GCATGC)$_2$ in 0.1 M NaCl, 0.05 M phosphate, pH=7.0, T=297 K.

carbon - 13 shift	G_1	C_2	A_3	T_4	G_5	C_6
C1'	85.0	86.8	85.7	85.4	85.0	87.2
C2'		40.0		39.5		42.0
C3'	80.2	78.5	80.7	76.9	79.8	71.6
C4'	89.8	86.5	88.3	85.6	87.7	86.8
C5'	65.0					
C2			155.4			
C5		98.9				97.7
C6		143.4		138.7		144.0
C8	139.4		142.5		138.7	
5CH_3				14.6		
coupling $^{1}J_{CH}$						
C1'-1'H	167	167	165	170	167	172
C3'-3'H	160	160	160	158	160	158
C4'-4'H	155	150	150	147	150	145
C2 - 2H			204			
C5 - 5H		170				172
C6 - 6H		178		182		176
C8 - 8H	216		214		210	
5CH_3				128		

a Chemical shifts are referenced to internal TSP (2,2,3,3-tetradeutero-3-(trimethylsilyl)-propionate, sodium salt) and are within ± 0.3 ppm.

b $^{1}J_{CH}$ in Hertz ± 6 Hertz

with the negative right hand component of a 2'CH cross peak of either G_1, A_3, or G_5. The C2' resonances of the latter nucleotides were not individually assigned, because of the very small chemical shift dispersion between 40.5 and 40.9 ppm. For C_6 both protons 2' and 2" have the same chemical shift (17). Only the 5'CH_2 resonances of G_1 were assigned, based on the well-separated, previously assigned proton resonance (17). All other C5' resonances could be located between 67.4 and 69.3 ppm (Fig.2B), but were not further distinguished because of the lack of sequence-specific ^{1}H NMR assignments (17). No coupling constants for 2'CH_2 or 5'CH_2 are included in Table 1. For the assigned resonances (Fig. 2B, 2C) $^{1}J_{CH}$ was estimated to be 143±5 Hz.

A first practical application of the ^{1}H-^{13}C correlations results from the larger dispersion of the ^{13}C chemical shifts, which enables unambiguous

^{1}H NMR assignments in situations where these would be difficult from ^{1}H NMR alone. For example, the ^{13}C resonances of the 4' and 5' carbon atoms are well separated, whereas the corresponding ^{1}H resonances 4'H, 5'H and 5"H exhibit strong overlap in the ^{1}H NMR spectrum. The combination of homonuclear ^{1}H COSY and heteronuclear [^{1}H,^{13}C]-COSY therefore allows an identification of the protons 4', 5' and 5". Further, the adenine C2 resonances have the lowest field ^{13}C shifts of all proton-bearing carbons and are well separated from the other resonances. It thus becomes straightforward to assign the adenine C2-protons, which can otherwise only be identified with relaxation time measurements, deuteration studies, or from NOE's to other protons in the duplex (19,24,25). Haasnoot and coworkers analyzed the ^{13}C NMR spectra of several oligoribonucleotides (12) and proposed a correlation between the ribose ^{13}C chemical shifts and parameters connected to conformational equilibria, e.g., the C2'endo-C3'endo equilibrium, and variations of the backbone torsion angles. For the presently studied hexanucleotide no detailed conformational data are available, so that the validity of these correlations cannot be further checked using this example. It is interesting with regard to future continuation of such studies, however, that the dispersion of the ^{13}C chemical shifts for the sugar carbons 3', 4', and 5' is approximately twice as large as for 1' and 2'.

[^{1}H,^{13}C]-COSY gives also information on the heteronuclear coupling constants $^1J_{CH}$ (Table 1). The couplings observed here in a double-stranded oligonucleotide are closely similar to those in the corresponding deoxynucleosides (20). The coupling constants $^1J_{CH}$ are much larger for the purine ring carbons (210±10 Hz) than for the pyrimidine ring carbons (175±8 Hz), so that by [^{1}H,^{13}C]-COSY the ring protons of pyrimidines and purines can be distinguished from the heteronuclear coupling constants $^1J_{CH}$. The coupling constants $^1J_{CH}$ of all 1' positions have about the same values (167±5 Hz), and are quite different from $^1J_{CH}$ of 3'C (160±5 Hz) or 4'C (150±5 Hz).

Overall, the experiment of Figs.1 and 2 clearly shows that [^{1}H,^{13}C]-COSY experiments with DNA fragments can be most useful at two different stages of NMR assignments. At the outset of an investigation, the ^{1}H-^{13}C correlations can help in the identification of the resonances from the different types of hydrogen atoms. Once sequence-specific ^{1}H NMR assignments have been obtained with the now well established procedures (e.g.,19,23-26), sequence-specific ^{13}C NMR assignments can then readily be obtained with these experiments.

In practice the [^{1}H,^{13}C]-COSY experiment can be optimized by adjustment

of the delay time to $\tau = 1/2(^1J_{CH})$. Rather than repeating the experiment with a different value of τ for each value of $^1J_{CH}$ which occurs in the molecule studied, a compromise for good efficiency may be found by selecting a value corresponding to an average coupling constant. For the present work with DNA the τ value of 3.6 ms corresponds to $^1J_{CH}$ of 140 Hz. The large deviation of the $^1J_{CH}$ values for the ring carbons of purines and pyrimidine from this average value explains the weaker intensity of the corresponding cross peaks when compared to the cross peaks involving ribose CH fragments (Figs. 1 and 2). It is also found that most cross peaks corresponding to CH_2 fragments have lower intensities than those corresponding to CH, as was also observed in the $[^1H,^{13}C]$-COSY spectrum of BPTI (11). As was discussed in connection with the cross peaks involving the 2' ^{13}C resonances (Fig.2C), one reason for this is that the doublets of the cross peaks C2'-2'H and C2'-2"H partially cancel each other, because the negative right hand component of one doublet comes close to, or coincides with the positive left hand component of the other doublet.

ACKNOWLEDGEMENTS

Financial support was obtained from the Swiss National Science Foundation (Project 3.198.85), the Swiss Cancer League (to W.L.), the Auckland Division of the Cancer Society of New Zealand and from the Medical Research Council of New Zealand. We thank Mrs. E.H. Hunziker and Mr. R. Marani for the careful preparation of the illustrations and the manuscript.

REFERENCES

1. Müller, L. (1979) J.Am.Chem.Soc. 101, 4481-4484.
2. Maudsley, A.A. and Ernst, R.R. (1977) Chem.Phys.Lett. 50, 368-372.
3. Bendall, M.R., Pegg, D.T. and Doddrell, D.M. (1983) J.Magn.Reson. 52, 81-117.
4. Bax, A., Griffey, R.H. and Hawkins, B.L. (1983) J.Am.Chem.Soc. 105, 7188-7190.
5. Bax, A., Griffey, R.H. and Hawkins, B.L. (1983) J.Magn.Reson. 55, 301-315.
6. Frey, M.H., Wagner, G., Vašák, M. Sørensen, O.W., Neuhaus, D., Wörgötter, E., Kägi, J.H.R., Ernst, R.R. and Wüthrich, K. (1985) J.Am.Chem.Soc. 107, 6847-6851.
7. Live, D.H., Armitage, I.M., Dalgarno, D.C. and Cowburn, D. (1985) J.Am. Chem.Soc. 107, 1775-1777.
8. Live, D.H., Davis, D.G., Agosta, W.C. and Cowburn, D. (1985) J.Am.Chem. Soc. 106, 6104-6105.
9. Otvos, J.D., Engeseth, H.R. and Wehrli, S. (1985) J.Magn.Reson. 61, 579-584.
10. Brühwiler, D. and Wagner, G. (1986) J.Magn.Reson., in press.
11. Wagner, G. and Brühwiler, D., (1986) Biochemistry, in press.

12. Lankhorst, P.P., Erkelens, C., Haasnoot, C.A.G. and Altona, C. (1983) Nucleic Acids Res. 11, 7215-7230.
13. Lankhorst, P.P., Haasnoot, C.A.G., Erkelens, C., Westerink, H.P., van der Marel, G.A., van Boom, J.H. and Altona, C. (1985) Nucleic Acids Res. 13, 927-942.
14. Borer, P.N., Zanatta, N., Holak, T.A., Levy, G.C., van Boom, J.H. and Wang, A.H.-J. (1984) J.Biomol.Struct.Dyn. 1, 1373-1386.
15. Shindo, H. (1981) Eur.J.Biochem. 120, 309-312.
16. Berman, E., Brown, S.C., James, T.L. and Shafer, R.H. (1985) Biochemistry 24, 6887-6893.
17. Frey, M.H., Leupin, W., Sørensen, O.W., Denny, W.A., Ernst, R.R. and Wüthrich, K. (1985) Biopolymers 24, 2371-2380.
18. Denny, W.A., Leupin, W. and Kearns, D.R. (1982) Helv.Chim.Acta 65, 2372-2393.
19. Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
20. Chang, C., Da Silva Gomes, J. and Byrn, S.R. (1983) J.Org.Chem. 48, 5151-5160.
21. Stothers, J.B. (1972) Carbon-13 NMR Spectroscopy, pp. 468-479, Academic Press, New York.
22. Rosenthal, S.N. and Fendler, J.H. (1976) Adv.Phys.Org.Chem. 13, 279-423.
23. Feigon, J., Leupin, W., Denny, W.A. and Kearns, D.R. (1983) Biochemistry 22, 5943-5951.
24. Scheek, R.M., Russo, N., Boelens, R., Kaptein, R. and van Boom, J.H. (1983) J.Am.Chem.Soc. 105, 2914-2916.
25. Hare, D.R., Wemmer, D.E., Chou, S.-H., Drobny, G. and Reid, B.R. (1983) J.Mol.Biol. 171, 319-336.
26. Weiss, M.A., Patel, D.J., Sauer, R.T. and Karplus, M. (1984) Proc.Nat. Acad.Sci.USA 81, 130-134.